specialbulletin från

PEDAGOGISK-PSYKOLOGISKA INSTITUTIONEN

LÄRARHÖGSKOLAN MALMÖ



# testkonstruktion och testdata

Bierschenk, B., Bierschenk, I. & Sternerup-Hansson, A.:

## ETT DATORPROGRAM FÖR SYNTAKTISK KODNING AV TITLAR TILL VETENSKAPLIGA SKRIFTER

Nr 36 Juli 1979

# ETT DATORPROGRAM FÖR SYNTAKTISK KODNING AV TITLAR TILL VETENSKAPLIGA SKRIFTER

Bernhard Bierschenk
Inger Bierschenk
Agneta Sternerup-Hansson

Bierschenk, B., Bierschenk, I. & Sternerup-Hansson, A.
Ett datorprogram för syntaktisk kodning av titlar till vetenskapliga skrifter. /A computer program for syntactic coding of titles to scientific documents./ Testkonstruktion och Testdata (Malmö, Sweden: Dept. of Educ. and Psychol. Research), Nr 36, 1979.

Användarorienterad kodning av kunskap som förmedlas via titlar till vetenskapliga skrifter kräver informationssystem som bygger på kognitiva modeller. I denna rapport presenteras ett datorprogram med kommentarer för en konceptuell kodning av titlar. Den kognitiva modell som styrt programutvecklingen beskrivs kortfattat. Programmets funktionssätt illustreras med några datorutskrifter från en databas som är representativ för svensk utbildningsforskning.

Nyckelord: Datalingvistik, datorprogram, informationsvetenskap, intermediära språk, kognition, rekognition, vetenskaplig dokumentation.

INNEHÅLL Sid

1. INLEDNING 3

2. EN KOGNITIONSPSYKOLOGISK MODELL 6

3. PROGRAMUTVECKLING 9

4. GENERERING AV MÖNSTER 21

5. GENERERING AV REGISTER 24

6. AVSLUTANDE DISKUSSION 38

7. REFERENSER 41

## 1. INLEDNING

I denna rapport redovisas kortfattat några steg i utvecklingen av ett datorbaserat informations- och dokumentationssystem. En utförlig redovisning av de teoretiska förutsättningarna för detta arbete finns i I. Bierschenk (1979).

Det datorprogram som presenteras har utvecklats inom projektet "Informations- och dokumentationsfrågor: Terminologi, datorbaserat system" som stöds av Skolöverstyrelsen. Projektet har som sitt huvudmål att utveckla ett system för spridning av information av intresse för utbildningssektorn. Med hjälp av titlar till vetenskapliga verk och de verk som citeras i dessa (referenserna) studeras inom projektet hur begreppsstrukturen ser ut inom utbildningsforskningen. För detta syfte har byggts upp experimentella databaser, nämligen som

1. innehåller alla titlar till vetenskapliga dokument som ett representativt stickprov av forskare inom utbildningsforskningen har producerat under en fyrtioårsperiod (1937-1977),
2. innehåller alla referenser som citerats i det stickprov av dokument som nämnts under punkt 1,
3. innehåller en koppling av referenserna via de enskilda dokumenten till ett omfattande intervjumaterial angående respektive forskares projektverksamhet som föreligger.

Att på detta sätt kunna kartlägga vilka erfarenheter (empirisk förankring) som ligger till grund för bildandet av beteendevetenskapliga begrepp och på vilket sätt de uppfattas av forskare som är verksamma inom utbildningssektorn skulle kunna bidra till utvecklingen av empiriskt förankrade informationssystem.

Att kunna utnyttja den information som finns i titlar till vetenskapliga skrifter förutsätter att det kan utvecklas program som tar hänsyn till strukturella förhållanden i språk. Detta kräver att regelsystem kan skapas som separerar dokumentbeskrivningar med avseende på deras

språkområden. Ett program för en automatisk separering har utvecklats (se I. Bierschenk, 1978).

Ett dokument kan beskrivas på olika nivåer. En rent bibliografisk beskrivning är enkel att såväl åstadkomma som automatiskt hantera. Men för återvinning av strukturerad information, dvs kunskap, måste dokumentbeskrivningen vara ett resultat av en innehållsanalys, vilket tillför informations- och dokumentationssystemet en kognitiv dimension. Kärnproblemet i utvecklingen av ett datorprogram för hantering av information är således att finna vägar för att representera kunskapen så att den kognitiva strukturen i vilken författaren meddelar den motsvaras av den struktur i vilken informationssökaren uppfattar den eller önskar den.

Det datorprogram som presenteras i denna rapport har utvecklats i syfte att kunna generera ett intermediärt språk, dvs ett språk som lämpar sig för en representation av vetenskaplig information. Med hjälp av programmet ska information kunna kodas användarorienterat. Under 1980-talet förväntas att många inom industrin och inom utbildningssektorn kommer att fråga efter informationstjänster som förutsätter denna typ av program, dvs program som kan ge "lösningar på problem".

Programmet utnyttjar abstrakta lingvistiska manifesteringar vid kodning av koncept och konceptuella relationer i titlar till vetenskapliga skrifter. Det genererar funktionellt relaterade enheter och inordnar dessa i ett antal olika dataregister.

Att analysera innehåll i en titel genom att ta de strukturella relationerna till hjälp har följande fördelar: (1) att val av nyckelord binds till ett regelsystem och (2) att titelförfattarens egen strukturering representeras.

Med utgångspunkt i den uttalade målsättningen att skapa en algoritm som gör information tillgänglig, så att den enskilde kan anpassa den till den kognitiva struktur som kännetecknar de föreställningar (modeller) som styr hans informationssökning, har problemområdet avgränsats till att gälla "MANAGEMENT OF KNOWLEDGE".

Med kunskap avses här strukturerad information. Att skapa överblick över all kunskap inom en viss disciplin är knappast realistiskt, särskilt som kunskapen ständigt utsätts för förändring. Det är därför programmets adaptiva egenskaper som återspeglar dess möjlighet att strukturera kunskap flexibelt.

Vetenskaplig rapportering har till syfte att kommunicera sådana kausala kopplingar som har kunnat etableras mellan händelser. Generaliseras detta påstående till en titel kan sägas att titeln till ett vetenskapligt dokument utgör ett abstraherat påstående om en forskningsprocess. Utvecklingen av en algoritm för en automatisk analys förutsätter således att processens enskilda komponenter måste kunna urskiljas för förståelsen av titeln.

## 2. EN KOGNITIONSPSYKOLOGISK MODELL

Språket är människans huvudsakliga medel för förmedlande av kunskap. En analys av sådan kunskap som förmedlas genom titlar till vetenskapliga skrifter förutsätter att en psykolingvistisk modell kan utvecklas. Det verkar nämligen naturligt att det i språket finns en konstruktion som gör det möjligt att avgränsa forskningsprocessens enskilda komponenter. Denna avgränsande och ordningsskapande funktion tycks prepositionerna ha.

En processmodell och en teori som bygger på en "Analys av Concept via Dator", som fått namnet ANACONDA, har utvecklats av Bierschenk & Bierschenk (1976).

Modellen bygger på en gestaltpsykologisk ansats, men skiljer sig också på många olika sätt från typiska gestaltpsykologiska tillvägagångssätt. Det som en gestaltpsykologisk modell har gemensamt med ANACONDA är att båda modellerna bygger på antagandet att det finns en mekanism som strukturerar data i enheter eller "helheter". Det som skiljer ANACONDA från en typisk gestaltpsykologisk modell är att ANACONDA inkorporerar erfarenheter och fenomenet uppmärksamhet som centrala komponenter. Denna skillnad blir särskilt tydlig vid operationaliseringen. ANACONDA nyttjar hypotesbildning och procedurer för hypotesprövning i syfte att studera symbolstrukturer. En typisk operationalisering av en gestaltpsykologisk modell skulle däremot innebära ett försök att studera mångfalden av parallellt uppträdande karakteristika som tillhör olika objekt (t ex titlar) och som ska särskiljas så att det framträder meningsfulla mönster.

ANACONDA är kontextorienterad och bygger på antagandet att den manifesta (språkliga) nivån kan utnyttjas för uppbyggande av ett schema för analys av latenta nivåer. Den utgår från antagandet att en mening består av en räcka mångtydiga ord eller symboler. Dessa implicerar i sin tur en komplex struktur av bakomliggande koncept. Genom en analys av koncept och relationer mellan koncept kan denna räcka transformeras till en struktur som kännetecknas av entydiga

"knutpunkter", som är kopplade med varandra genom ett explicit angivet regelsystem. Begreppet "knut" refererar till den lexikala betydelsen av ett ord, medan "regelsystemet" anger de relationer som specificerar och sammanbinder ord (t ex modifierare, kvantifierare, klassificerare, subjekt-verb-objekt-kopplingar).

Med ANACONDA kan för en viss bestämd sats vid en viss given tidpunkt prediceras relativt entydiga koncept och konceptrelationer. Vid utprövningen av ANACONDA (se B. Bierschenk, 1977) har kunnat visas att konceptuell information kan kodas och att sådan information kan adresseras rätt.

På den manifesta nivån centreras delar av en mening runt verbet, vilket ger verbet med dess olika partiklar en nyckelfunktion. I detta ändras heller ingenting när ANACONDA tillämpas på avsnitt eller större sammanhängande texter. Temat hos ett avsnitt kan ses som ett "makroverb", dvs som en aktion eller som en karakterisering av något som gör något. Konceptualiseringen av ett textstycke ses med utgångspunkt i ANACONDA som en syntes av en verbstruktur.

Verbets nyckelfunktion består i att det uttrycker ett tillstånd och att det anger vilken funktion som ska tilldelas $S_1$ och $S_2$ inom Substantiv$_1$-verb-Substantiv$_2$-paradigmet. Ur kognitionspsykologisk synvinkel innebär detta dock endast en övergångsfas. Verbets transienta funktion leder över till en intermediär nivå som karakteriseras av Agent-aktion-Objekt-paradigmet. Genom AaO-paradigmet avgränsas de komponenter som utgör ett naturligt sammanhang.

Resultatet av en syntaktisk analys karakteriseras av dess transienta egenskap, dvs det bibehålls endast så länge som transformationsprocessen kräver det. Resultatet av framgångsrika transformeringar är en växande kunskapsstruktur.

En kunskapsrepresentation byggd på en modell (schema) är rationell och spar lagringsutrymme. Ett schema förutsätter abstraktioner och nyttjar information från konceptuella relationer. En modell byggd på schemaantagandet söker svara på vad som ska aktiveras för att ett system ska kunna ge önskat svar.

Utmärkande för ett schema är att dess komponenter alltid föreligger. Komponenterna måste emellertid förses med variabler, vilket innebär att ett schema operationaliseras. Vilka värden som de enskilda variablerna kommer att anta beror däremot på den kontext inom vilken ett schema aktiveras. Kontext definieras här som de omständigheter med vilka individen är förtrogen.

Skapande av kunskap måste ses som en konstruktiv aktivitet, som består i ett uppbyggande av lämpliga scheman. Kunskap skapas genom vetenskaplig aktivitet. Vetenskap kan enligt Bunge (1967) inte betraktas som ett bestämbart objekt eller en bestämbar uppsättning av problem, utan måste betraktas som ett sätt att angripa problem (metod och mål). Den avgränsning av vetenskaplig aktivitet som behövs för en analys av titlar till vetenskapliga skrifter kan skapas genom Problem-metod-Mål-paradigmet. Detta är ett paradigm av högre ordning än $S_1$ v $S_2$- och AaO-paradigmet och antas äga giltighet oavsett plats och tid. PmM-paradigmet ska kunna ge information om den kontextuella begränsning som titelförfattare anger som villkor för problemlösning. Kunskap kan således också uppfattas som förtrogenhet med ett problemrum ("problem space"). Detta begrepp har lånats från Newell & Simon (1972). Den kontext inom vilken ett empiriskt fenomen analyseras kan ses som gräns för informationssökarens tolkningsmöjligheter. Den syntaktiska strukturen, som styr urvalet av termer och sammansättningen av de transienta titelschemana, antas innehålla lingvistiska indikatorer på den kunskap och de kontextuella begränsningar som titelförfattaren vill förmedla till det vetenskapliga samfundet.

## 3. PROGRAMUTVECKLING

PmM-paradigmet har till uppgift att återspegla ordningen inom en titel. Uppgiften att skapa ordning inom titlar innefattar en bestämning av vad en titel är. I detta syfte har utvecklats ett regelsystem (se I. Bierschenk, 1979). Genom avgränsningsreglerna anges en synlig gräns för början och slut av en konceptualisering. Två typer av regler utgör sedan grunden för den konceptuella kodningen, nämligen stoppregler och struktureringsregler.

Jämfört med den intermediära nivån finns det på den manifesta nivån relativt få valmöjligheter när det gäller att förmedla kunskap. Däremot finns det stor frihet att nyttja "kontext" som en villkorlig ("optional") variabel.

En förutsättning för programutveckling är därför att PmM-paradigmet förses med flexibla egenskaper, dvs första uppgiften var att tilldela dem "default"-variabler (default = frånvaro) som platshållare. Med hjälp av dessa kan PmM-paradigmet adapteras så att det passar till vissa specifika förhållanden. Relationerna mellan olika specifika förhållanden antas bli återspeglade av de relationer som existerar mellan olika "markörer". Markörernas huvudsakliga uppgift är att ange till vilken komponent i modellen en viss bestämd enhet (ett datum) ska tillordnas. Programmet bygger på antagandet att prepositioner i en titel pekar mot bestämda typer av koncept. Prepositioner antas fungera som markörer. I detta sammanhang är deras position i förhållande till varandra av stor betydelse. De markörer som är tillordnade modellens huvudkomponenter är "av", "med" och "för". Dessa för "handlingen" framåt och anger således riktningen i paradigmet.

Prepositioner som avgränsar huvudkomponenterna är placerade mellan dessa och i viss ordning, vilket medför att de koncept som de pekar mot sorteras in under närmast föregående huvudpreposition.

Huvudkomponenternas markörer antas vara bärande. Denna grundstruktur realiserar intentioner och kan ses som en representation av abstraherade erfarenheter. En sådan grundstruktur kan sedan förses med olika många komponenter som underordnas huvudkompo-

nenterna. Medan grundkomponenterna alltid representerar ett och samma fenomen kan de underordnade komponenterna bete sig mera flexibelt. Markörerna till dessa betraktas som extensionella, eftersom de kan förändras genom ytterligare uppdelningar eller genom att de kan anta olika parametervärden. På så sätt kan de underordnade komponenterna lätt anpassa sig till specifika uppsättningar av vetenskapliga titlar.

Intentionella prepositioners "data" (markerade med huvudpreposition enligt ruta 1) tilldelas en kodsiffra som slutar på "0". Det tillhörande extensionella datum (markerat med övrig preposition) tilldelas en annan andrasiffra, där ordningen börjar med annan siffra än noll. Systemet upprätthålls på det sättet att ordningsföljden rättas efter varje ny huvudkod, vilket innebär att strukturen binds algoritmiskt.

Prepositioner i titlar kan ha varierande kontexter, vilka exakt ska kunna bestämmas, dvs prepositionernas "längd" måste definieras. Det finns nämligen ett antal flerordsuttryck som kan jämställas med prepositioner. Prepositioner kan också ingå i fasta fraser, där de inte har någon pekande funktion. Prepositioner kan dessutom kombineras med varandra, antingen i direkt anslutning eller länkade med hjälp av en konjunktion. Preposition och konjunktion kan även bilda ett par.

Specificering av sådana fall förhindrar "feltolkning". Vilka specificeringar programmet arbetar med visas genom det lexikon som presenteras i ruta 1.

Eftersom programmet bygger på att huvudbegrepp särskiljs från underordnade, har reglerna utformats så att också längre titlar kan få en korrekt kodning. En längre titel kan nämligen innehålla flera fall av samma preposition. Oavsett prepositionens status i modellen kan en titel endast innehålla en huvudkomponent av samma typ. De regler som ligger till grund för programutvecklingen presenteras i ruta 2.

Baserat på de regler som presenterats i ruta 2 har skrivits ett datorprogram i ASC II-FORTRAN. Programmet har utprövats på UNIVAC 1108. Det kodar en text syntaktiskt genom

Ruta 1. Lexikon vid kodning av titlar

Prepositioner är:*

| | | |
|---|---|---|
| i | som | mot |
| av | mellan | enligt |
| för | från | genom |
| på | rörande | inför |
| till | under | åt |
| med | hos | ur |
| om | kring | över |
| inom | bland | angående |
| vid | efter | |

Flerordsprepositioner är: *

| | Funktion |
|---|---|
| i anslutning till | = i |
| i samband med | = i |
| i relation till | = i |
| i fråga om | = i |
| med hjälp av | = med |
| med speciellt avseende på | = i |
| med särskild hänsyn till | = i |
| med särskild anknytning till | = i |
| med särskild inriktning på | = i |
| med tonvikt lagd vid | = i |
| med tonvikt lagd på | = i |
| kombinerad med | = i |
| jämförd med | = i |

Flerordsfraser är:

typer av
slag av
grad av
former av
ramen för
samvariation med
samverkan med
ett perspektiv av
synpunkter på
exempel på
redogörelse för
aspekter på

Konjunktioner är:

och
eller
samt
resp
,

Huvudprepositioner är:

| | | Funktion |
|---|---|---|
| av | om | = av |
| med | angående | = av |
| för | rörande | = av |
| | över | = av |

* Operationell definition

1. att leta reda på givna strängar (ord),
2. att nyttja ordens relation till varandra och
3. att tilldela texten sifferkoder för olika syntagmer (ett eller flera ord).

Programmet har skrivits för kodning av ett speciellt textmaterial, nämligen en uppsättning titlar till vetenskapliga verk som författats av ett stickprov forskare som är verksamma inom svensk utbildningsforskning. Detta innebär att programmet i första hand har skrivits för en speciell användning. På sidorna 13-20 återges en kommenterad version av programmet.

Ruta 2. Reglerna för identifiering av koncept och syntaktisk kodning av titlar till vetenskapliga skrifter

| | |
|---|---|
| Regel 1: | Kontrollera flerordsprepositioner |
| Regel 2: | Kontrollera flerordsfraser |
| Regel 3: | Preposition och konjunktion inom ( ) eller mellan / verkar ej |
| Regel 4: | Preposition som första ord verkar ej |
| Regel 5: | Om konjunktion sammanbinder två prepositioner = konjunktionen och andra prepositionen verkar ej |
| Regel 6: | Om preposition följs av konjunktion = prepositionen verkar ej |
| Regel 7: | Om konjunktion följs av preposition = prepositionen verkar ej |
| Regel 8: | Om preposition följs av preposition = första prepositionen verkar ej |
| Regel 9: | Om med eller för följs av sig själv = andra förekomsten blir underordnad |
| Regel 10: | Endast en av av, om, rörande, angående eller över kan vara huvudpreposition i en sats. Den första förekomsten blir huvudpreposition. Sats avgränsas av konjunktion eller kommatecken |
| Regel 11: | Preposition som sista ord verkar ej |
| Regel 12: | Huvudregel för inledande preposition |

| | | |
|---|---|---|
| 40 | av<br>om<br>rörande<br>angående<br>över | 30 |
| 40 | för | 70 |
| 40 | med | 80 |
| 30 | ej huvud-<br>preposition | 33 |

Endast första preposition i sats bestämmer framåt.

Regel 13: Huvudregel för ej inledande preposition

| | | | | | | |
|---|---|---|---|---|---|---|
| ___ | av<br>: | 30 | ej huvud-<br>preposition | 33 | ej huvud-<br>preposition | 34 |
| ___ | för | 70 | ej huvud-<br>preposition | 73 | ej huvud-<br>preposition | 74 |
| ___ | med | 80 | ej huvud-<br>preposition | 83 | ej huvud-<br>preposition | 84 |
| ___ | | | ej huvud-<br>preposition | 33 | av<br>: | 34 |

Regel 14: Konjunktion sammanbinder uttryck av samma typ

```
000001   000          SUBROUTINE KODORD(MENING,NANTBO,KOD,ORD)
000002   000    C   ********************************************************************
000003   000    C   ***                                                              ***
000004   000    C   *** SUBROUTINEN KODORD LETAR I EN MENING REDA PÅ PREP , KONJ     ***
000005   000    C   *** OCH ANDRA ORD SOM INGÅR I FLERORDSFRASER OCH FLERORDSKONJ    ***
000006   000    C   *** OCH MÄRKER UT DEM MED ETT TAL BEREOENDE PÅ VILKEN FUNKTION   ***
000007   000    C   *** DE SKALL HA I SUBROUTINEN BESTAM. DESSA TAL LAGRAS I KOD.    ***
000008   000    C   ***                                                              ***
000009   000    C   *** MENING(I)=ETT ORD I AKTUELL MENING (IN).                     ***
000010   000    C   *** NANTBO(I)=ANTALET BOKSTAVER I ORDET MENING(I) (IN).          ***
000011   000    C   *** ORD=ANTAL ORD I AKTUELL MENING (IN).                         ***
000012   000    C   ***                                                              ***
000013   000    C   ***     KOD(I)=0  OM ORD NR I EJ ÄR EN OPERATOR                  ***
000014   000    C   ***  1<=KOD(I)<=5 OM ORD NR I ÄR EN INTENTIONELL PREP            ***
000015   000    C   *** 11<=KOD(I)<=20 OM ORD NR I AR EN EXTENSIONELL PREP           ***
000016   000    C   ***     KOD(I)=30  OM ORD NR I AR EN KONJ                        ***
000017   000    C   *** 40<=KOD(I)<=99 OM ORD NR I AR ETT OVRIGT ORD I OPERATORLISTA***
000018   000    C   ***     KOD(I)=100 OM ORD NR I ÄR EN PARENTES                    ***
000019   000    C   ***                                                              ***
000020   000    C   *** LISTA = EN LISTA ÖVER OPERATORER ORDNAD ENLIGT ORDLANGD.     ***
000021   000    C   *** PEKA EN PEKARE FÖR SÖKNINGEN I LISTA.                        ***
000022   000    C   *** PEKA(K) ANGER VAR SÖKNINGEN SKALL BÖRJA FÖR ETT ORD SOM      ***
000023   000    C   *** HAR LÄNGDEN K.                                               ***
000024   000    C   *** NLIST(K)= ANTALET ORD SOM HAR LANGDEN K.                     ***
000025   000    C   ***                                                              ***
000026   000    C   ********************************************************************
000027   000          IMPLICIT INTEGER (A-Z)
000028   000          DIMENSION NANTBO(100),KOD(100),NR(64),PEKA(12),NLIST(12)
000029   000          CHARACTER MENING*25(100)
000030   000          CHARACTER LISTA*12(64)
000031   000          DATA NR/30,100,100,100,
000032   000         *11,20,1,12,4,20,2,21,3,15,20,20,20,30,55,54,13,14,20, 5,30,50,
000033   000         *30,50,20,20,20,20,20,20,30,50,51,42,43,20,20,50,47,5,
000034   000         *61,53,41,59,60,40,45,5,46,52,44,61,40,47,30,56,57,48,58,52/
000035   000          DATA LISTA/
000036   000         *',          ','(          ',')          ',
000037   000         *'/          ','I          ',
000038   000         *'UR         ','AV         ','PÅ         ',
000039   000         *'OM         ','ÅT         ','FÖR        ','ATT        ',
000040   000         *'MED        ','VID        ','SOM        ',
000041   000         *'HOS        ','MOT        ','OCH        ','ETT        ',
000042   000         *'LAGD       ','TILL       ','INOM       ','FRÅN       ',
000043   000         *'ÖVER       ','SAMT       ','SLAG       ','RESP       ',
000044   000         *'GRAD       ','UNDER      ','KRING      ',
000045   000         *'BLAND      ','EFTER      ','GENOM      ',
000046   000         *'INFÖR      ','ELLER      ','TYPER      ',
000047   000         *'RAMEN      ','FRÅGA      ','HJALP      ',
000048   000         *'MELLAN     ','ENLIGT     ','FORMER     ',
000049   000         *'HÄNSYN     ','RÖRANDE    ','JÄMFORD    ','TONVIKT    ',
000050   000         *'SAMBAND    ','EXEMPEL    ','ASPEKTER   ',
000051   000         *'RELATION   ','AVSEENDE   ','ANGÅENDE   ',
000052   000         *'SÄRSKILD   ','SAMVERKAN  ','SPECIELLT  ',
000053   000         *'KOMBINERAD ','ANSLUTNING ','ANKNYTNING ','RESPEKTIVE ',
000054   000         *'PERSPEKTIV ','SYNPUNKTER ','INRIKTNING ','REDOGORELSE ',
000055   000         *'SAMVARIATION'/
000056   000          DATA PEKA/1,6,11,20,29,40,44,49,54,56,63,64/
000057   000          DATA NLIST/5,5,9,9,11,4,5,5,2,7,1,1/
000058   000          DO 100 I=1,100
000059   000    100   KOD(I)=0
000060   000          DO 110 I=1,ORD
000061   000          J=NANTBO(I)
000062   000          IF (J.GT.12)  GOTO 110
000063   000          IF (NLIST(J).EQ.0)  GOTO 110
000064   000          I1=PEKA(J)
000065   000          I2=PEKA(J)+NLIST(J)-1
000066   000          DO 120 K=I1,I2
000067   000    120   IF (LISTA(K).EQ.SUBSTR(MENING(I),1,12)) GOTO 130
000068   000          GOTO 110
000069   000    130   KOD(I)=NR(K)
000070   000    110   CONTINUE
000071   000          RETURN
000072   000          END
```

```
000001   000   C   ************************************************************************
000002   000   C   ***                                                                  ***
000003   000   C   *** PROGRAMMET SYNTAXKODAR EN FIL MED REFERENSER . SKRIVER UT        ***
000004   000   C   *** FILEN MED SYNTAXMARKERINGAR PÅ RADSKRIVAREN . DESSUTOM           ***
000005   000   C   *** SKRIVER PROGRAMMET UT VARJE MENINGSMÖNSTER MED DESS              ***
000006   000   C   *** IDENTIFIKATIONSNR PÅ EN FIL OCH PÅ EN ANNAN FIL SKRIVES          ***
000007   000   C   *** VARJE SYNTAX UT MED TILLHÖRANDE TEXT OCH IDENTIFIKATIONSNR .***
000008   000   C   ***                                                                  ***
000009   000   C   *** INDATA FIL.                                                      ***
000010   000   C   ***     FIL NR 8 .                                                   ***
000011   000   C   ***             KOL 1-8 IDENTIFIKATIONSNR                            ***
000012   000   C   ***             KOL 9  KORT NR ( KORT NR 2 INNEHÅLLER REF TITEL )    ***
000013   000   C   ***             KOL 10 ANTAL KORT AV SAMMA TYP                       ***
000014   000   C   ***             KOL 11 SEQNR INOM KORT NR                            ***
000015   000   C   ***             KOL 21-80 TEXT                                       ***
000016   000   C   ***                                                                  ***
000017   000   C   *** UTDATA FILER.                                                    ***
000018   000   C   ***     FIL 10  MÖNSTER FILEN.                                       ***
000019   000   C   ***             KOL 1-8 IDENTIFIKATIONSNR                            ***
000020   000   C   ***             KOL 9-128 MÖNSTER                                    ***
000021   000   C   ***                                                                  ***
000022   000   C   ***     FIL 11  SYNTAX FILEN.                                        ***
000023   000   C   ***             KOL 1-8 IDENTIFIKATIONSNR                            ***
000024   000   C   ***             KOL 10-11 SYNTAXKOD                                  ***
000025   000   C   ***             KOL 13-132 TEXT                                      ***
000026   000   C   ***                                                                  ***
000027   000   C   ************************************************************************
000028   000          IMPLICIT INTEGER (A-Z)
000029   000          DIMENSION KOD(100),NANTBO(100),TAB(100),SYNTAX(100),SYNT(100)
000030   000          CHARACTER MENING*25(100)
000031   000          CHARACTER STRING*1(1000)
000032   000          CHARACTER ST*2
000033   000          CHARACTER IDNR*8
000034   000   50     IF (EOF.EQ.1) GOTO 999
000035   000   51     CALL LAS(STRING,NANT,IDNR,EOF)                     É INLÄSNING AV DATA
000036   000          IF (EOF.EQ.1.AND.NANT.EQ.0)  GOTO 999              É TEST AV FILSLUT
000037   000          NG2=1
000038   000          TECKEN=0
000039   000          ORD=1
000040   000          DO 100 I=1,NANT
000041   000   C   ***
000042   000   C   ***  TEST AV MENINGSSLUT
000043   000   C   ***  EN MENING AVSLUTAS MED ? ! . : ELLER - OMGIVET AV BLANKA
000044   000   C   ***
000045   000          IF (STRING(I).EQ.'?'.OR.
000046   000        *    STRING(I).EQ.'!'.OR.
000047   000        *    STRING(I).EQ.':'.OR.
000048   000        *    STRING(I).EQ.'.') GOTO 120
000049   000          IF (I.EQ.1)  GOTO 106
000050   000          IF (STRING(I-1).EQ.' '.AND.
000051   000        *    STRING(I).EQ.'-'.AND.
000052   000        *    STRING(I+1).EQ.' ') GOTO 120
000053   000   105    IF (STRING(I).EQ.' ')  GOTO 110
000054   000   106    TECKEN=TECKEN+1
000055   000          SUBSTR(MENING(ORD),TECKEN,1)=STRING(I)
000056   000          GOTO 100
000057   000   110    IF (TECKEN.EQ.0)  GOTO 100
000058   000          NANTBO(ORD)=TECKEN
000059   000          LLL=TECKEN+1
000060   000          DO 115 LL=LLL,25
000061   000   115    SUBSTR(MENING(ORD),LL,1)=' '
000062   000          TAB(ORD)=I-NANTBO(ORD)
```

```
000063     000          ORD=ORD+1
000064     000          TECKEN=0
000065     000          NANTBO(ORD)=0
000066     000          GOTO 100
000067     000    C     ***
000068     000    C     ***  NY MENING
000069     000    C     ***
000070     000    120   NG1=NG2
000071     000          NG2=I+1
000072     000          IF (NANTBO(ORD).EQ.0) ORD=ORD-1
000073     000    C     ***
000074     000    C     ***  MÄRK UT MENINGENS OPERATORER .
000075     000    C     ***
000076     000          CALL KODORD(MENING,NANTBO,KOD,ORD)
000077     000    C     ***
000078     000    C     ***  SYNTAXKODA EN MENING
000079     000    C     ***
000080     000          CALL BESTAM(KOD,SYNTAX,ORD)
000081     000          IF (STRING(NG1).EQ.' ') NG1=NG1+1
000082     000    C    ***
000083     000    C    ***  UTSKRIFT AV EN MENING
000084     000    C    ***
000085     000          WRITE(6,150)  IDNR,(STRING(LL),LL=NG1,NG2)
000086     000    150   FORMAT(/1X,A8,2X,(1X,120A1))
000087     000          DO 1190 LL=1,ORD
000088     000          IF (LL.EQ.1) L1=1
000089     000          IF (LL.EQ.1.AND.ORD.NE.1) GOTO 1190
000090     000          IF (SYNTAX(LL).EQ.0.AND.LL.NE.ORD) GOTO 1190
000091     000          LL1=TAB(L1)
000092     000          IF  (LL.EQ.ORD) LL2=NG2-2
000093     000          IF  (LL.NE.ORD) LL2=TAB(LL)-1
000094     000    C    ***
000095     000    C    *** UTSKRIFT AV EN SYNTAGM MED TILLHÖRANDE TEXT
000096     000    C    ***
000097     000          WRITE(11,2000) IDNR,SYNTAX(L1),(STRING(L2),L2=LL1,LL2)
000098     000          L1=LL
000099     000    2000  FORMAT(A8,1X,I2,1X,120A1)
000100     000    1190  CONTINUE
000101     000          DO 160 LL=NG1,NG2
000102     000    160   STRING(LL)=' '
000103     000          DO 170 LL=1,ORD
000104     000          IF (SYNTAX(LL).EQ.0)  GOTO 170
000105     000          LLL=TAB(LL)
000106     000          ENCODE(230,ST) SYNTAX(LL)
000107     000    230   FORMAT(I2)
000108     000          STRING(LLL)=SUBSTR(ST,1,1)
000109     000          STRING(LLL+1)=SUBSTR(ST,2,1)
000110     000    170   CONTINUE
000111     000    C    ***
000112     000    C    *** UTSKRIFT AV SYNTAGMMARKERINGAR PÅ RADSKRIVAREN
000113     000    C    ***
000114     000          WRITE(6,180) IDNR,(STRING(LL),LL=NG1,NG2)
000115     000    180   FORMAT(1X,A8,2X,(1X,120A1))
000116     000          L1=0
000117     000          DO 1170 LL=1,ORD
000118     000          IF (SYNTAX(LL).EQ.0) GOTO 1170
000119     000          L1=L1+1
000120     000          SYNT(L1)=SYNTAX(LL)
000121     000    1170  CONTINUE
000122     000    C    ***
000123     000    C    ***  UTSKRIFT AV MÖNSTERFIL
000124     000    C    ***
000125     000          WRITE(10,1180) IDNR,(SYNT(IJ),IJ=1,L1)
000126     000    1180  FORMAT(A8,40(1X,I2))
000127     000          ORD=1
000128     000          TECKEN=0
000129     000    100   CONTINUE
000130     000          GOTO 50
000131     000    C    ***
000132     000    C    ***  FILSLUT
000133     000    C    ***
000134     000    999   WRITE(6,200)
000135     000          ENDFILE 10
000136     000          ENDFILE 11
000137     000    200   FORMAT( '1  ALLT KLART' )
000138     000          END
```

```
000001    000          SUBROUTINE LAS(STRING,NANT,IDNR,EOF)
000002    000     C    ************************************************************************
000003    000     C    ***                                                                  ***
000004    000     C    ***  SUBROUTINEN LAS LÄSER EN TITEL I VERKBASEN SOM ÄR FIL NR 8      ***
000005    000     C    ***  REDIGERAR OM TITELN SÅ ATT ALLA ORD OCH SPEC.-TECKEN OMGES      ***
000006    000     C    ***  AV BLANKA SAMT KONTROLLERAR ATT TITELN AVSLUTAS MED SKILJE-     ***
000007    000     C    ***  TECKEN .                                                        ***
000008    000     C    ***                                                                  ***
000009    000     C    ***  STRING INNEHÅLLER EN TITEL EFTER ANROP AV LAS                   ***
000010    000     C    ***  NANT ÄR ANTALET TECKEN I STRING                                 ***
000011    000     C    ***  IDNR ÄR VERKETS IDENTIFIKATIONSNR                               ***
000012    000     C    ***  EOF=1 OM FILSLUT FÖR ÖVRIGT 0                                   ***
000013    000     C    ***                                                                  ***
000014    000     C    ************************************************************************
000015    000          IMPLICIT INTEGER(A-Z)
000016    000          CHARACTER STRING*1(1000)
000017    000          CHARACTER ST*1(1000)
000018    000          CHARACTER IDNR*8
000019    000          NANT=0
000020    000     50   READ(8,100,END=999) IDNR,KNR,NANTK,(ST(I),I=1,60)
000021    000     100  FORMAT(A8,2I1,10X,60A1)
000022    000          IF (KNR.NE.2) GOTO 50
000023    000          I1=1
000024    000          I2=60
000025    000          ST(61)=' '
000026    000          IF (NANTK.EQ.1)  GOTO 125
000027    000          DO 120 I=2,NANTK
000028    000          I1=I1+61
000029    000          I2=I2+61
000030    000          READ(8,101,END=999)  (ST(IJ),IJ=I1,I2)
000031    000     101  FORMAT(20X,60A1)
000032    000          ST(I2+1)=' '
000033    000     120  CONTINUE
000034    000     125  J=0
000035    000          DO 130 I=1,I2
000036    000          N=BITS(ST(I),1,9)
000037    000          IF (N.EQ.32) GOTO 160
000038    000          IF ((N.GE.65.AND.N.LE.93).OR.
000039    000         *   (N.EQ.45).OR.
000040    000         *   (N.GE.48.AND.N.LE.57))  GOTO 140
000041    000     C    ***
000042    000     C    ***  ST(I) INNEHÅLLER SPEC TECKEN
000043    000     C    ***
000044    000          IF (STRING(J).EQ.' '.OR.J.EQ.0) J=J-1
000045    000          IF (J.EQ.-1) GOTO 139
000046    000          STRING(J+1)=' '
000047    000     139  STRING(J+2)=ST(I)
000048    000          STRING(J+3)=' '
000049    000          J=J+3
000050    000          GOTO 130
000051    000     C    ***
000052    000     C    ***  ST(I) INNEHÅLLER BOKSTAV ELLER BINDESTRECK
000053    000     C    ***
000054    000     140   J=J+1
000055    000          STRING(J)=ST(I)
000056    000          GOTO 130
000057    000     C    ***
000058    000     C    ***  ST(I) INNEHÅLLER BLANKT
000059    000     C    ***
000060    000     160  IF (J.EQ.0)  GOTO 130
000061    000          IF (STRING(J).EQ.' ')  GOTO 130
000062    000          J=J+1
000063    000          STRING(J)=' '
000064    000     130  CONTINUE
000065    000     131  IF (STRING(J).NE.' ') GOTO 170
000066    000          J=J-1
000067    000          GO TO 131
000068    000     170  CONTINUE
000069    000     C    ***
000070    000     C    ***  KONTROLLERA ATT SISTA TECKNET AR ETT MENINGSSKILJANDE TECKEN
000071    000     C    ***
000072    000          IF (STRING(J).EQ.'?'.OR.
000073    000         *STRING(J).EQ.'!'.OR.STRING(J).EQ.':'.OR.
000074    000         *STRING(J).EQ.'.')  GOTO 173
000075    000          IF (STRING(J).EQ.'-'.AND.STRING(J-1).EQ.' ') GOTO 173
000076    000          STRING(J+1)=' '
000077    000          STRING(J+2)='.'
000078    000          J=J+2
000079    000     173  NANT=J
000080    000          RETURN
000081    000     C    ***
000082    000     C    ***  FILSLUT
000083    000     C    ***
000084    000     999   EOF=1
000085    000           RETURN
000086    000           END
```

```
000001    000         SUBROUTINE BESTAM(KOD,SYNTAX,NANT)
000002    000    C    *****************************************************************
000003    000    C    ***                                                           ***
000004    000    C    *** SUBROUTINEN BESTAM SYNTAXKODAR EN MENING EFTER DET ATT    ***
000005    000    C    *** KODORD KLASSIFICERAT DE INGÅENDE ORDEN .                  ***
000006    000    C    ***                                                           ***
000007    000    C    *** KOD (IN) . DETTA FÅR MAN FRÅN SUBROUTINEN KODORD .        ***
000008    000    C    *** NANT (IN) . ANTALET ORD I MENINGEN SOM SKALL SYNTAXKODAS. ***
000009    000    C    *** SYNTAX (UT) . MENINGENS SYNTAGMER .                       ***
000010    000    C    ***                                                           ***
000011    000    C    *****************************************************************
000012    000         INTEGER KOD(NANT),SYNTAX(NANT)
000013    000         INTEGER FLERPR(4,10),FLERFR(3,8),SY
000014    000         DATA FLERPR/3,44,45,12,           É MED SPECIELLT AVSEENDE PÅ
000015    000        *            3,46,47,13,           É MED SÄRSKILD HANSYN TILL
000016    000    C                                        MED SÄRSKILD ANKNYTNING TILL
000017    000        *            3,46,48,12,           É MED SÄRSKILD INRIKTNING PÅ
000018    000        *            3,53,54,15,           É MED TONVIKT LAGD VID
000019    000        *            3,53,54,12,           É MED TONVIKT LAGD PÅ
000020    000        *            3,43,1,0,             É MED HJALP AV
000021    000        *            11,40,13,0,           É I ANSLUTNING TILL
000022    000        *            11,41,3,0,            É I SAMBAND MED
000023    000        *            11,42,4,0,            É I FRÅGA OM
000024    000        *            61,3,0,0/             É KOMBINERAD MED , JÄMFÖRD MED
000025    000         DATA FLERFR/55,56,1,              É ETT PERSPEKTIV AV
000026    000        *            51,2,0,               É RAMEN FÖR
000027    000        *            52,3,0,               É SAMVARIATION MED
000028    000    C                                        SAMVERKAN MED
000029    000        *            57,12,0,              É SYNPUNKTER PÅ
000030    000        *            50,1,0,               É TYPER AV , SLAG AV
000031    000    C                                        GRAD AV , FORMER AV
000032    000        *            58,2,0,               É REDOGÖRELSE FÖR
000033    000        *            59,12,0,              É EXEMPEL PÅ
000034    000        *            60,12,0/              É ASPEKTER PÅ
000035    000         DO 90 I=1,NANT
000036    000    90   SYNTAX(I)=0
000037    000    C    ***
000038    000    C    *** FLERORDSPREPOSITIONER .                                 REGEL 1
000039    000    C    ***
000040    000         IF (NANT.LE.3) GOTO 130           É 4 ORD LÅNGA
000041    000         NT=NANT-3
000042    000         DO 100 I=1,NT
000043    000         DO 110 J=1,5
000044    000         IF (KOD(I).EQ.FLERPR(1,J).AND.
000045    000        *    KOD(I+1).EQ.FLERPR(2,J).AND.
000046    000        *    KOD(I+2).EQ.FLERPR(3,J).AND.
000047    000        *    KOD(I+3).EQ.FLERPR(4,J)) GOTO 120
000048    000    110  CONTINUE
000049    000         GOTO 100
000050    000    C    ***
000051    000    C    *** FLERORDSPREP MOTSVARAS AV EN EXTENSIONELL PREP .        REGEL 1
000052    000    C    ***
000053    000    120  KOD(I+1)=0
000054    000         KOD(I+2)=0
000055    000         KOD(I+3)=0
000056    000         KOD(I)=20
000057    000    100  CONTINUE
000058    000    130  IF (NANT.LE.2) GOTO 180           É 3 ORD LÅNGA
000059    000         NT=NANT-2
000060    000         DO 140 I=1,NT
000061    000         DO 150 J=6,9
000062    000         IF (KOD(I).EQ.FLERPR(1,J).AND.
```

```
000063    000          *    KOD(I+1).EQ.FLERPR(2,J).AND.
000064    000          *    KOD(I+2).EQ.FLERPR(3,J)) GOTO 160
000065    000    150   CONTINUE
000066    000          GOTO 140
000067    000    160   KOD(I+1)=0
000068    000          KOD(I+2)=0
000069    000    140   CONTINUE
000070    000    180   CONTINUE
000071    000          IF (NANT.LE.1)  GOTO 1180           É 2 ORD LÅNGA
000072    000          NT=NANT-1
000073    000          DO 1140 I=1,NT
000074    000          DO 1150 J=10,10
000075    000          IF (KOD(I).EQ.FLERPR(1,J).AND.
000076    000          *    KOD(I+1).EQ.FLERPR(2,J)) GOTO 1160
000077    000    1150  CONTINUE
000078    000          GOTO 1140
000079    000    1160  KOD(I+1)=0
000080    000          KOD(I)=20
000081    000    1140  CONTINUE
000082    000    1180  CONTINUE
000083    000    C   ***
000084    000    C   ***  FLERORDSFRASER .                                          REGEL 2
000085    000    C   ***
000086    000          IF (NANT.LE.2) GOTO 175             É 3 ORD LÅNGA
000087    000          NT=NANT-2
000088    000          DO 170 I=1,NT
000089    000          DO 171 J=1,1
000090    000          IF (KOD(I).EQ.FLERFR(1,J).AND.
000091    000          *    KOD(I+1).EQ.FLERFR(2,J).AND.
000092    000          *    KOD(I+2).EQ.FLERFR(3,J)) GOTO 172
000093    000    171   CONTINUE
000094    000          GOTO 170
000095    000    172   KOD(I)=0
000096    000          KOD(I+1)=0
000097    000          KOD(I+2)=0
000098    000    170   CONTINUE
000099    000    175   IF (NANT.LE.1)  GOTO 220            É 2 ORD LÅNGA
000100    000          NT=NANT-1
000101    000          DO 190 I=1,NT
000102    000          DO 200 J=2,3
000103    000          IF (KOD(I).EQ.FLERFR(1,J).AND.
000104    000          *    KOD(I+1).EQ.FLERFR(2,J))  GOTO 210
000105    000    200   CONTINUE
000106    000          GOTO 190
000107    000    210   KOD(I)=0
000108    000          KOD(I+1)=0
000109    000    190   CONTINUE
000110    000    220   CONTINUE
000111    000    C   ***
000112    000    C   ***  PREP OCH KONJ INOM PARENTES VERKAR EJ .                  REGEL 3
000113    000    C   ***
000114    000          DO 400 I=1,NANT
000115    000          IF (KOD(I).NE.100) GOTO 400
000116    000          KOD(I)=0
000117    000          DO 410 J=I,NANT
000118    000          IF (KOD(J).EQ.100) GOTO 420
000119    000          KOD(J)=0
000120    000    410   CONTINUE
000121    000    420   KOD(J)=0
000122    000    400   CONTINUE
000123    000    C   ***
000124    000    C   ***
000125    000    C   ***  PREPOSITIONER FÖRST I MENING VERKAR EJ .                 REGEL 4
000126    000    C   ***
```

```
000127     000          DO 230 I=1,NANT
000128     000          IF (KOD(I).EQ.0.OR.KOD(I).GE.40)  GOTO  240
000129     000    230   CONTINUE
000130     000    240   DO 250 J=1,I
000131     000    250   KOD(J)=0
000132     000    C     ***
000133     000    C     ***  KONJ SOM SAMMANBINDER TVÅ PREP VERKAR EJ .
000134     000    C     ***  ANDRA PREP VERKAR EJ .                                REGEL 5
000135     000    C     ***
000136     000          IF (NANT.LE.2) GOTO 270
000137     000          NT=NANT-2
000138     000          DO 260 I=1,NT
000139     000          IF (KOD(I).GE.1.AND.KOD(I).LE.20.AND.KOD(I+1).EQ.30.AND.
000140     000         *    KOD(I+2).GE.1.AND.KOD(I+2).LE.20)  GOTO 280
000141     000          GOTO 260
000142     000    280   KOD(I+1)=0
000143     000          KOD(I+2)=0
000144     000    260   CONTINUE
000145     000    270   CONTINUE
000146     000    C     ***
000147     000    C     *** OM PREP FÖLJS AV KONJ DÅ VERKAR EJ PREP .              REGEL 6
000148     000    C     *** OM KONJ FÖLJS AV PREP DÅ VERKAR EJ PREP .              REGEL 7
000149     000    C     *** OM PREP FÖLJS AV PREP DÅ VERKAR EJ FÖRSTA PREP .       REGEL 8
000150     000    C     ***
000151     000          IF (NANT.LE.1) GOTO 277
000152     000          NT=NANT-1
000153     000          DO 275 I=1,NT
000154     000          IF (KOD(I).GE.1.AND.KOD(I).LE.20.AND.KOD(I+1).EQ.30) KOD(I)=0
000155     000          IF (KOD(I+1).GE.1.AND.KOD(I+1).LE.20.AND.KOD(I).EQ.30) KOD(I+1)=0
000156     000          IF (KOD(I+1).GE.1.AND.KOD(I+1).LE.20.AND.KOD(I).GE.1.AND.
000157     000         *KOD(I).LE.20) KOD(I)=0
000158     000    275   CONTINUE
000159     000    277   CONTINUE
000160     000    C     ***
000161     000    C     *** OM MED ELLER FÖR FÖLJS AV SIG SJÄLV DÅ BLIR
000162     000    C     *** ANDRA FOREKOMSTEN EN EXTENSIONELL PREP .               REGEL 9
000163     000    C     ***
000164     000          N1=0
000165     000          DO 300 I=1,NANT
000166     000          IF (KOD(I).EQ.0)  GOTO 300
000167     000          N2=KOD(I)
000168     000          IF (N1.EQ.KOD(I).AND.(KOD(I).GE.2.AND.KOD(I).LE.3)) KOD(I)=20
000169     000          N1=N2
000170     000    300   CONTINUE
000171     000    C     ***
000172     000    C     *** ENDAST EN AV PREP ANGÅENDE,RÖRANDE,OM,ÖVER ELLER AV KAN VARA
000173     000    C     *** INTENTIONELL PREP I EN SATS OCH FÅR EJ FOREGÅS AV EN
000174     000    C     *** EXTENSIONELL PREP. SATS AVGRÄNSAS AV , ELLER KONJ.     REGEL 10
000175     000    C     ***
000176     000          DO 311 I=1,NANT
000177     000          IF (KOD(I).EQ.0.OR.KOD(I).EQ.2.OR.KOD(I).EQ.3)  GOTO 311
000178     000          IF (KOD(I).GT.20)  GOTO 311
000179     000          II=I+1
000180     000          DO 312 J=II,NANT
000181     000          IF (KOD(J).EQ.1.OR.KOD(J).EQ.4.OR.KOD(J).EQ.5) KOD(J)=20
000182     000          IF (KOD(J).EQ.30) GOTO 313
000183     000    312   CONTINUE
000184     000    313   CONTINUE
000185     000    311   CONTINUE
000186     000    C     ***
000187     000    C     *** PREP I SLUTET AV MENING VERKAR EJ                      REGEL 11
000188     000    C     ***
000189     000          IF (KOD(NANT).GE.1.AND.KOD(NANT).LE.20) KOD(NANT)=0
000190     000    C     ***
```

```
000191        000   C    ***  REDIGERINGAR OCH OMDEFINITIONER AVKLARADE ENLIGT REGLERNA
000192        000   C    ***  1 TILL 11 .
000193        000   C    ***
000194        000   C    ***  HUVUDREGLER .
000195        000   C    ***
000196        000        NFORST=0
000197        000        SY=0
000198        000        DO 320 I=1,NANT
000199        000        IF (KOD(I).EQ.0.OR.KOD(I).GE.21)  GOTO 320
000200        000        IF (NFORST.EQ.1)  GOTO 330
000201        000        NFORST=1
000202        000   C    ***
000203        000   C    *** SYNTAGMKODNING AV INLEDANDE DELEN AV EN SATS .          REGEL 12
000204        000   C    ***
000205        000        IF (KOD(I).GE.1.AND.KOD(I).LE.5)  SYNTAX(1)=40
000206        000        IF (KOD(I).GE.6.AND.KOD(I).LE.20)  SYNTAX(1)=30
000207        000        SY=SYNTAX(1)
000208        000   C    ***
000209        000   C    *** SYNTAGMKODNING AV EFTERFÖLJANDE DELAR AV EN SATS .      REGEL 13
000210        000   C    ***
000211        000   330  IF (KOD(I).EQ.1.OR.KOD(I).EQ.4.OR.KOD(I).EQ.5) SYNTAX(I)=30
000212        000        IF (KOD(I).EQ.3)   SYNTAX(I)=80
000213        000        IF (KOD(I).EQ.2)   SYNTAX(I)=70
000214        000        IF (KOD(I).GE.1.AND.KOD(I).LE.5) SY=SYNTAX(I)
000215        000        IF (KOD(I).GE.1.AND.KOD(I).LE.5)  GOTO  320
000216        000        IF (SY.EQ.30.OR.SY.EQ.80.OR.SY.EQ.70)  SY=SY+2
000217        000        SY=SY+1
000218        000        SYNTAX(I)=SY
000219        000   320  CONTINUE
000220        000        IF (NFORST.EQ.0)  SYNTAX(1)=30
000221        000   C    ***
000222        000   C    ***  KONJ SAMMANBINDER ELEMENT .                            REGEL 14:1
000223        000   C    ***
000224        000        DO 360 I=1,NANT
000225        000        IF (KOD(I).NE.30) GOTO 360
000226        000        KONJI=I
000227        000        L=KONJI+1
000228        000        IF (L.GT.NANT) GOTO 360
000229        000        DO 370 J=L,NANT
000230        000   370  IF (KOD(J).NE.0) GOTO 380
000231        000        GOTO 360
000232        000   380  IF (KOD(J).GE.1.AND.KOD(J).LE.5)   SYNTAX(KONJI)=40
000233        000   360  CONTINUE
000234        000   C    ***
000235        000   C    *** KONJ SAMMANBINDER TVÅ UTTRYCK AV SAMMA TYP .            REGEL 14:2
000236        000   C    ***
000237        000        SY=SYNTAX(1)
000238        000        DO 340 I=1,NANT
000239        000        IF (SYNTAX(I).NE.0)   SY=SYNTAX(I)
000240        000        IF (KOD(I).EQ.30.AND.SYNTAX(I).EQ.0)   SYNTAX(I)=SY
000241        000   340  CONTINUE
000242        000        RETURN
000243        000        END
```

## 4. GENERERING AV MÖNSTER

I det följande ska redovisas vilka mönster som kan urskiljas ur titlarna. Enligt analysmodellen kan konceptualiseringarna vara mer eller mindre utbyggda. I det mest utbyggda fallet represen-teras de av komponenterna 30 (Problem) + 40 (Metod) + (80) (Instru-ment) + 70 (Mål) med eventuella bestämningar (annan andrasiffra än 0). Ett mönster är således en strukturell representation av en konceptualisering, vilket implicerar att vissa mönster är möjliga, vanliga, ovanliga eller omöjliga. Dessutom finns det en bunden restriktion i systemet bl a genom de underordnade kodernas nummerföljd. I illustrerativt syfte presenteras i ruta 3 utskriften av de mönster som kännetecknar strukturen i titlarna hos försöks-person 24.

Ruta 3. Ett exempel på mönster i titlar

```
24001   30 30
24002   30
24002   40 30
24003   30
24003   30
24004   40 30 33
24005   30 30 33
24006   30
24006   40 30
24007   40 30
24008   30 33
24008   30
24012   30 33
24020   30 33 34 34
24021   40 70
24023   30 33
24023   30
24024   30 33
24024   30
24025   30
```

För förklaring, se ovan.

Ruta 3 visar att endast Problem-komponenten (30) och Metod-komponenten (40) förekommer initialt. Vidare framgår att Metod endast följs av huvudkomponent, främst 30.

En analys av mönstren presenteras i I. Bierschenk (1979, kap 5). Analysen visar att några generella drag kan urskiljas. Av möjliga mönster har inte många aktiverats. De mest sannolika mönstren består av en ensam Problemkomponent. Därefter följer en konnektion av två Problemkoncept eller en extensionell avgränsning. Ju mera sammansatta mönstren är desto mera sällan förekommer de.

Ytterligare analyser (I. Bierschenk, 1979, kap 5.3) visar att titlarnas strukturella variation samvarierar med den komplexitet som kännetecknar dokumenttyp. Ur informationsvetenskaplig synvinkel skulle detta resultat kunna läggas till grund för en differentiering och stegvis återvinning av information relaterad till dokumenttyp.

Vid kodningen antogs att ett koncept kan anta vilken lingvistisk representation som helst. Vidare antogs att den semantiska strukturen (baserad på AaO-paradigmet) är av transient natur. Med dessa utgångspunkter är alltså den kognitiva strukturen varken identisk med den syntaktiska eller den semantiska. PmM-schemat bestämmer således inte realiteten utan endast vad som kan accepteras från realiteten och hur den kan struktureras. En välfungerande återvinningsmekanism kan därför bara delvis bygga på syntaktiska och semantiska regler. Kan den inte nyttja information från PmM-paradigmet (se ruta 4), kan den inte heller leda till annat än föga relevanta och ganska oreliabla återvinningsresultat. Många lingvistiska objekt (ord eller fraser) har nämligen multipla syften och bör därför beskrivas snarare funktionellt än fysiskt. Detta innebär att de enheter som genereras genom programmet och ordnas i enskilda register bör kännetecknas av "funktionell likhet". Att enheterna fyller samma funktion innebär emellertid inte att de också nödvändigtvis måste vara homogena i andra avseenden än vad funktionen gäller. Ruta 4 exemplifierar en utskrift efter automatisk kodning av en försökspersons (24) titlar. De kodade enheterna utgör basen för registerbildning, vilken diskuteras i nästa kapitel.

Ruta 4. Exempel på syntaktisk kodning av titlar

```
24001     KOVARIANSANALYS OCH KVASI-EXPERIMENT .
24001     30                  30

24002     LORD ' S PARADOX -
24002     30

24002     OM ANALYS AV ICKE-RANDOMISERADE DATA .
24002     40    30

24003     RELATIV PRESTATION .
24003     30

24003     EN RECENSION .
24003     30

24004     TVÅ NOTISER OM MÄTNING AV FÖRÄNDRING .
24004     40         30         33

24005     ÅLDERS- OCH TIDSEFFEKTER I KOGNITIV UTVECKLING .
24005     30          30           33

24006     FÖRELÄSNING CONTRA ANDRA UNDERVISNINGSMETODER .
24006     30

24006     RAPPORT OM EN LITTERATURGRANSKNING .
24006     40      30

24007     MÄTNING AV MENTAL PRESTATIONSUTVECKLING .
24007     40      30

24008     STUDENTERNAS ANPASSNING TILL STUDIERNA .
24008     30                      33

24008     UNIVERSITETETS PEDAGOGIK .
24008     30

24012     STUDENTERNAS ANPASSNING TILL STUDIERNA .
24012     30                      33

24020     NÅGRA JÄMFÖRELSER MELLAN ELEVER I GYMNASIUM OCH FACKSKOLAN .
24020     30              33                34           34

24021     STANDARDPROVSRESULTAT FÖR DIALEKTTALANDE ELEVER .
24021     40                    70

24023     ELEVER I OBSERVATIONSKLINIK .
24023     30       33

24023     UNDERSÖKNING 1 .
24023     30

24024     ELEVER I OBSERVATIONSKLINIK .
24024     30       33

24024     UNDERSÖKNING II .
24024     30

24025     AR EVALUERING FORSKNING ?
24025     30
```

[illegible] = annat språk

## 5. GENERERING AV REGISTER

Registrens innehåll och funktion utgör den direkta grunden för utveckling och utprövning av både terminologi och återvinningsmekanism. Detta avsnitt har därför till syfte att förmedla en uppfattning om vad registren i huvudsak innehåller och hur de kan fungera. I ruta 5 presenteras enheterna med registerkod från de titlar som producerats av försöksperson 24.

Ruta 5. Exempel på innehåll i registren

```
24001   30 KOVARIANSANALYS
24001   30 OCH KVASI-EXPERIMENT
24002   30 LORD ' S PARADOX
24002   40 OM ANALYS
24002   30 AV ICKE-RANDOMISERADE DATA
24003   30 RELATIV PRESTATION
24003   30 EN RECENSION
24004   40 TVÅ NOTISER
24004   30 OM MÄTNING
24004   33 AV FÖRÄNDRING
24005   30 ÅLDERS-
24005   30 OCH TIDSEFFEKTER
24005   33 I KOGNITIV UTVECKLING
24006   30 FÖRELÄSNING CONTRA ANDRA UNDERVISNINGSMETODER
24006   40 RAPPORT
24006   30 OM EN LITTERATURGRANSKNING
24007   40 MÄTNING
24007   30 AV MENTAL PRESTATIONSUTVECKLING
24008   30 STUDENTERNAS ANPASSNING
24008   33 TILL STUDIERNA
24008   30 UNIVERSITETETS PEDAGOGIK
24012   30 STUDENTERNAS ANPASSNING
24012   33 TILL STUDIERNA
24020   30 NÅGRA JÄMFÖRELSER
24020   33 MELLAN ELEVER
24020   34 I GYMNASIUM
24020   34 OCH FACKSKOLAN
24021   40 STANDARDPROVSRESULTAT
24021   70 FÖR DIALEKTTALANDE ELEVER
24023   30 ELEVER
24023   33 I OBSERVATIONSKLINIK
24023   30 UNDERSÖKNING 1
24024   30 ELEVER
24024   33 I OBSERVATIONSKLINIK
24024   30 UNDERSÖKNING II
24025   30 ÄR EVALUERING FORSKNING
```

Att finna den rätta modellen för en reproduktion av den kognitiva strukturen i en titel och den information som titelförfattaren avser att förmedla är inte någon lätt sak. En välfungerande återvinningsmekanism borde därför bygga på en kombination av flera olika tekniker, t ex (1) symbolisk representation och logiska operatorer som arbetar med symboler, (2) numeriska analyser, t ex klustringstekniker och (3) partiell mönsterigenkänning.

Sådana naturliga datauppsättningar som redovisas i ruta 5 kännetecknas i regel av brist på tillräckligt distinkta kodningsenheter. Denna brist kan upphävas genom klassbildning. Det krävs alltså någon slags abstraktion. Vad som bör undvikas är emellertid en inflexibel (inklusionsorienterad) klassificering av enheterna.

Det finns knappast någon uppsättning egenskaper som simultant kan användas för att beskriva alla enheter i de bildade registren. De enheter som redovisas i ruta 5 har kodats med utgångspunkt i en symbolisk representation och operatorer som arbetar med symboler. Igenkänningen av elementen i varje enhet har skett med hjälp av markörer som angivit datatyp (för illustration, se ruta 4). Uppdelningen av de igenkända enheterna i klasser inom respektive register kräver däremot andra tekniker, t ex klustringstekniker eller tilldelning av deskriptorer.

En vokabulärstudie av samtliga register har genomförts (drygt 2100 genererade enheter) varvid några olika aspekter (fasetter) kan beskrivas. Ett första försök att gruppera detta stora antal enheter genomfördes med hjälp av klusteranalystekniker.

Klusteranalys valdes som metod eftersom den möjliggör en uppdelning av enheterna i oberoende kluster eller överlappande homogena klasser. Syftet var alltså att producera en operationell klassificering. Denna klassificering skulle sedan vara till hjälp vid en optimering av informationsåtervinningsprocesser (t ex en förkortning av sökprocessen eller vid identifiering av vissa informationstyper). I syfte att studera hållbarheten i klusteranalysresultaten används två olika klustringsalgoritmer, nämligen

1. en klustringsalgoritm som söker täta ("minimum varians") kluster och
2. en klustringsalgoritm som söker etablera "naturliga" kluster som inte nödvändigtvis behöver vara täta.

Leder en tillämpning av båda förfaringssätten till samma klustrings-resultat, kan de erhållna klustren betraktas som tillförlitliga. Det datorprogram som har använts är CLUSTAN (se Wishart, 1975).

Resultaten av försöken att etablera kluster som innehåller några få enheter som kan ses som typiska för respektive kluster är samstämmiga oavsett vilken algoritm som nyttjas. Det kan fastslås att detta material har en mycket låg struktureringsgrad. Oavsett vilket kriterium som tillämpas visar det sig vara svårt att finna meningsfulla kluster baserat på forskares sätt att använda koncept i titlar till vetenskapliga skrifter. I ruta 6 presenteras i illustrativt syfte några kluster ur en analys som bygger på den första klustrings-algoritmen.

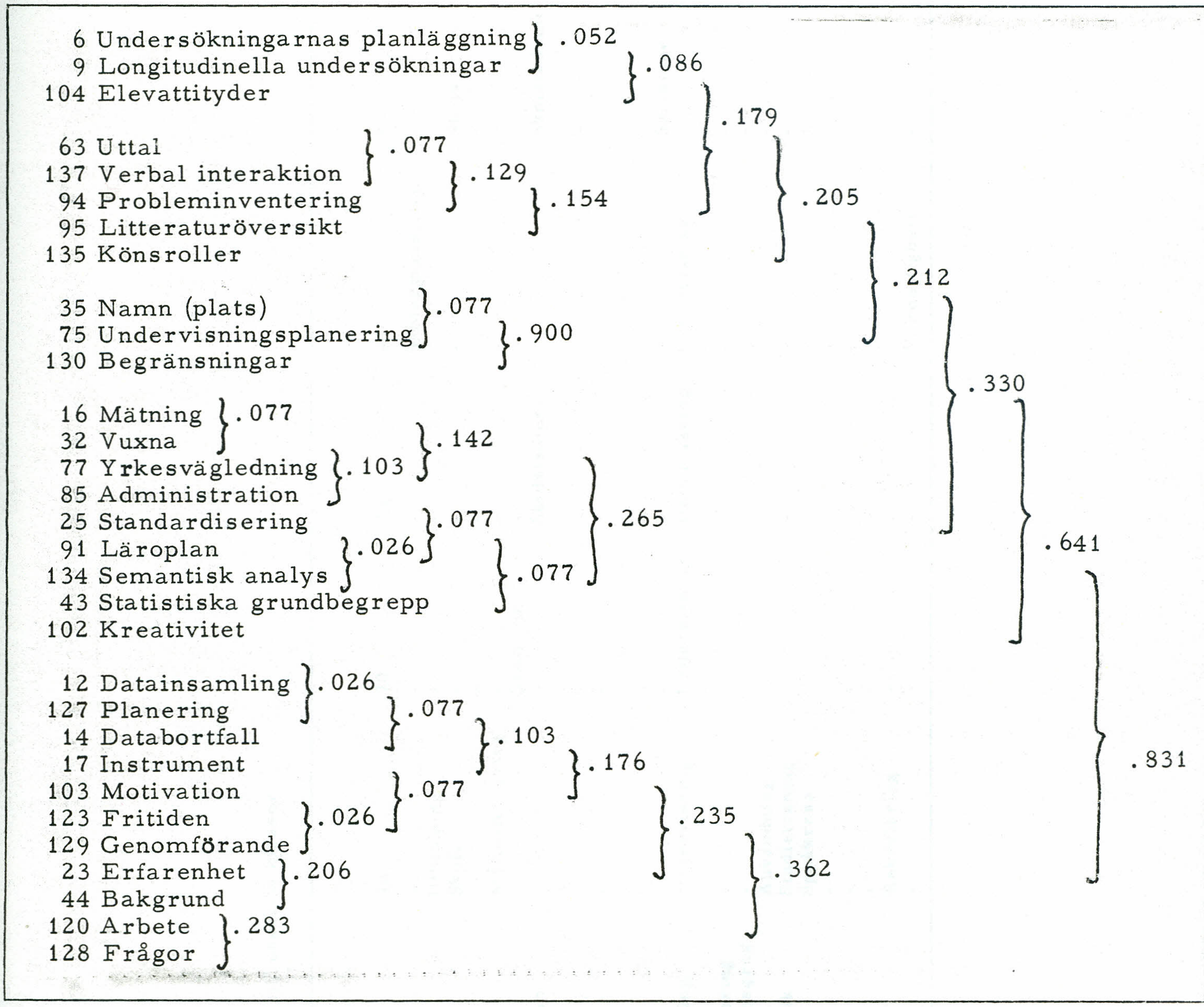
Som framgår ur ruta 6 är det svårt att finna någon lämplig "cut off". Dessutom är det svårt att se de gemensamma egenskaper som simultant är tillämpliga på alla enheter som bildar klustret.

Eftersom uppdelningen av enheterna med hjälp av klusteranalys-tekniker inte visade sig vara någon framkomlig väg, återstår bara en väg, nämligen tilldelning av deskriptorer.

Samtliga enheter skrevs på kort och sorterades sedan register för register i olika högar. En enhet betraktas lika en annan när

1. den består av samma element oavsett stavningsvarianter, t ex SÖ, skolöverstyrelsen,
2. den består av samma element med olika ändelser, t ex elev, elever,
3. den består av olika kvalificerande element, t ex amerikansk pedagogik, pedagogik,
4. den består av olika element med samma funktionella innebörd, t ex vetenskaplig belysning, pedagogisk synpunkt.

Ruta 6. Några kluster ur en klusteranalys som bygger på en minimering av variansen




Med utgångspunkt i dessa kriterier skapades 357 olika högar eller klasser som skulle benämnas. För att få en överblick över registrens likheter respektive olikheter tilldelades så långt som möjligt samma deskriptorer till en enskild klass. Principen illustreras i ruta 7.

Ruta 7. Tilldelning av deskriptorer till enskilda klasser

Huvuddeskriptor: Skolorganisation

| Register 30 | 33 | 40 | 70 | 73 | 80 | 83 |
|---|---|---|---|---|---|---|
| - | - | - | - | - | Skolorganisation | - |
| - | - | Utbildning | - | - | - | - |
| Skola | Skola | Skola | - | - | - | Skola |
| Mål | - | | - | - | - | - |
| Målsökning | - | Målformulering | - | - | - | - |
| - | Skolform | - | Skolform | - | - | - |
| Skolstadium | Skolstadium | - | - | Skolstadium | - | Skolstadium |
| - | Årskurs | - | - | - | - | - |
| Klassen | Klassen | - | - | - | - | - |
| - | Klinik | - | - | - | - | - |
| Befattning | Befattning | - | - | - | - | - |
| Didaktik | - | - | - | - | - | - |
| Skolämne | Skolämne | - | - | - | - | Skolämne |
| Undervisning | Undervisning | Undervisning | Undervisning | Undervisning | Undervisning | - |
| - | Lektion | - | - | - | - | - |
| Differentiering | Differentiering | - | - | - | - | - |
| Material-metodsystem | Konkretionsgrad | - | - | - | - | - |
| Studier | Studier | Användning | - | - | - | - |
| Studieteknik | Studieteknik | Studieträning | - | - | - | - |
| Språkfärdighet | | Språkresa | - | - | - | - |
| Lästräning | Läsning | - | - | - | - | - |
| Rättstavning | - | - | - | - | - | - |
| Inskolning | - | - | - | - | - | - |
| Rekrytering | - | Rekrytering | - | - | - | - |
| - | - | - | - | - | Valmöjligheter | - |

Delfasetten "Skola" inom huvudfasetten "Skolorganisation" i registren 30, 40 och 70 illustreras i ruta 8.

Ruta 8. Fasett "Skola" inom huvudfasett "Skolorganisation" ur registren 30, 40 och 70

| Register 30 | Register 40 | Register 70 |
|---|---|---|
| Skola* | Skola | - |
| Skolans övergripande mål | | |
| Skolans målsättning* | | |
| Skolans socialisation | | |
| Skolans utveckling* | | |
| Skolans kris* | | |
| Skolans sociologi | | |
| Skolnivå* | | |
| Skolsegregation* | | |

* trunkerade enheter

Ruta 8 visar vilka olika termer, utgående från "skola", som återfinns i respektive register. Termerna ska kunna fungera som sökord, vilket betyder att en matchning av dessa strängar mot "naturliga" titelbegrepp ska kunna ske. Trunkeringen (avgränsningen med hjälp av asterisk) anger gränser för den identitet som krävs vid en sådan matchning. Asterisken efter skola* anger alltså att ändelser tillåts förekomma. I vissa fall antas inga ändelser förekomma eller det tillåts inga tecken efter strängens slut.

Denna uppsättning trunkerade termer har till uppgift att representera varianter av samma term inom en fasett. Empiriska utprövningar får visa vilka trunkeringsprinciper som är optimala. Detta beror i sin tur på om och i så fall vilka kriterier för "likhet" som bör anges inom respektive fasett. Trunkeringen skol* skulle t ex endast ange en term i fasetten och betrakta alla andra som likvärdiga för återvinning.

En sammanfattning av huvudfasetterna och antalet delfasetter inom svensk utbildningsforskning presenteras i tabell 1. Tabellen visar att antalet underkategorier är mest frekvent inom register 30 (problem) och 33 (explicita avgränsningar). De huvudfasetter som krävt det flesta

Tabell 1. Fasetter och antal underfasetter inom svensk utbildningsforskning: Fördelningar ur en representativ experimentbas

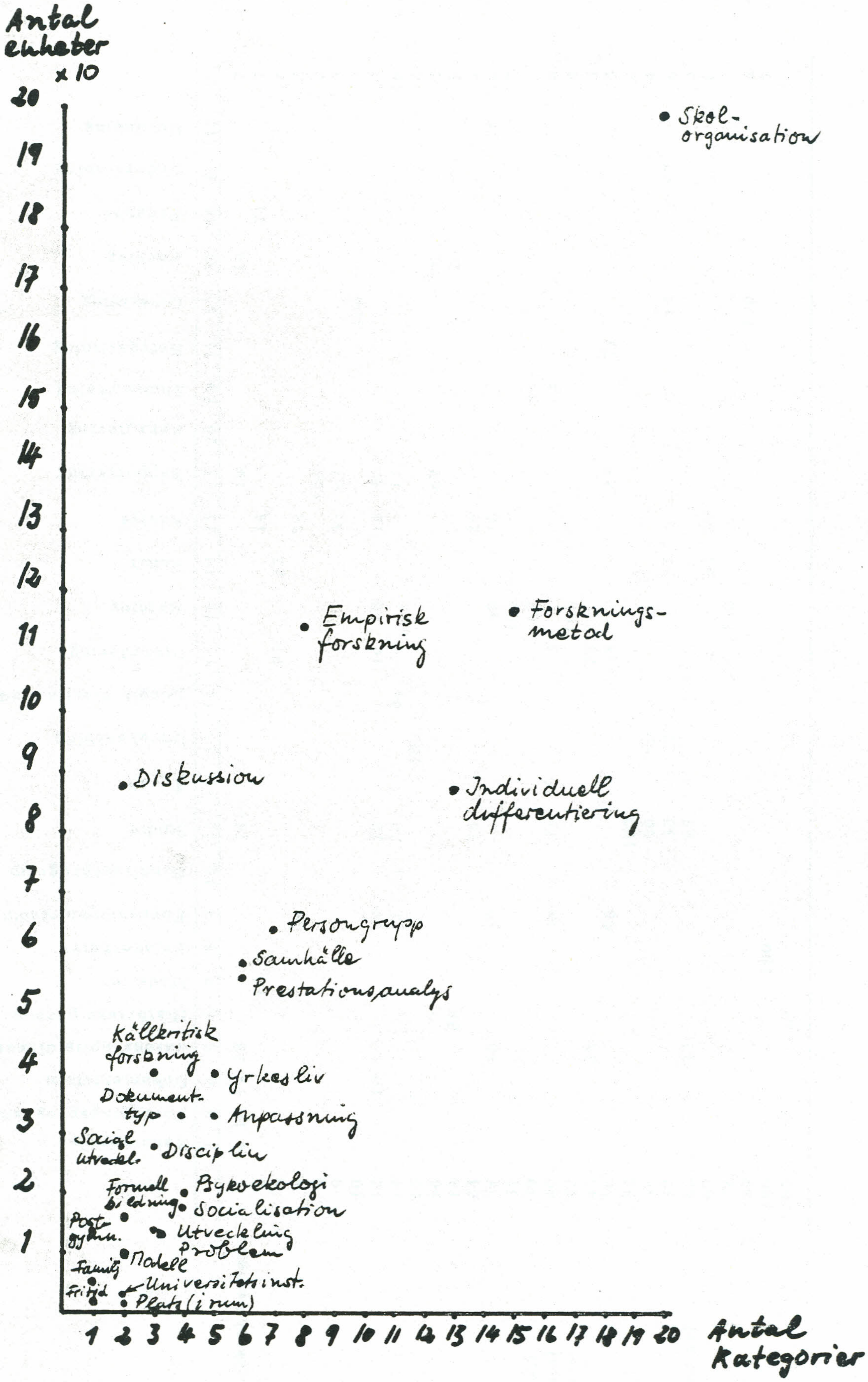
| Huvudfasetter | Antal Kod 30 | 33+... | 40 | 70 | 73+... | 80 | 83+... | Σ |
|---|---|---|---|---|---|---|---|---|
| Disciplin | 3 | 1 | 1 | | 1 | | | 6 |
| Diskussion | 2 | 5 | 2 | | | | | 9 |
| Modell | 2 | 2 | 1 | | | | | 5 |
| Problem | 3 | | | | 1 | | | 4 |
| Källkritisk forskning | 3 | 2 | 1 | | | | | 6 |
| Empirisk forskning | 8 | 4 | 8 | 1 | | | | 21 |
| Forskningsmetodik | 15 | 10 | 13 | 2 | 1 | 1 | 1 | 43 |
| Persongrupp | 7 | 6 | | 3 | 3 | 1 | 1 | 21 |
| Samhälle | 6 | 7 | | 2 | 1 | | | 16 |
| Yrkesliv | 5 | 8 | | 2 | 3 | | | 18 |
| Universitetsinstitution | 2 | | | | | | | 2 |
| Formell bildning | 2 | | | | | | | 2 |
| Socialisation | 4 | 5 | | | | 1 | 1 | 11 |
| Social utveckling | 2 | 1 | | 1 | | | | 4 |
| Anpassning | 5 | 3 | 2 | | | 1 | 1 | 12 |
| Familj | 1 | | | | | | | 1 |
| Fritid | 1 | | | | | | | 1 |
| Skolorganisation | 20 | 16 | 8 | 2 | 1 | 3 | 3 | 53 |
| Postgymnasial utbildning | 1 | 2 | | 1 | | | | 4 |
| Individuell differentiering | 13 | 7 | | 1 | 1 | | | 22 |
| Prestationsanalys | 6 | 8 | 4 | | 1 | | | 19 |
| Beteende | | 4 | | | | | | 4 |
| Psykoekologi | 4 | 1 | | | | | | 5 |
| Utveckling | 3 | 1 | 1 | 1 | | | | 6 |
| Dokumenttyp | 4 | 4 | | | | | | 8 |
| Plats (i rum) | 2 | 4 | | | 1 | | | 7 |
| Plats (i tid) | | 5 | | 1 | | | | 6 |
| Σ | 124 | 106 | 41 | 58 | 14 | 7 | 7 | 357 |
| Σ enheter | 1107 | 566 | 255 | 74 | 35 | 42 | 18 | 2092 |

antalet undergrupper är "Forskningsmetodik", "Empirisk forskning", "Skolorganisation" och "Individuell differentiering". Detta kan ses som ett tecken på vad utbildningsforskningen huvudsakligen sysslat med under en 40-årsperiod. Den huvudfasett som benämns "Diskussion" har däremot i förhållande till antalet enheter endast några få underkategorier. Differentieringsgraden inom problemkomponenten åskådliggörs grafiskt i figur 1. Ett annat allmänt drag är att de register som ligger längre bort från register 30 innehåller alltmera specificerad information vilket kan utläsas ur förhållandet mellan antalet enheter per register och antalet nödvändiga delfasetter. Innehållet och strukturen inom enheterna i registren ska här inte närmare kommenteras. Det kommer att studeras i samband med tesaurkonstruktionen.

Som framgår ur figur 1 kan ytterligare några grupperingar urskiljas, nämligen "Persongrupp", "Samhälle" och "Prestationsanalys". De problemområden som tycks vara relativt nya eller speciella är "Formell bildning", "Postgymnasial utbildning", "Familj", "Fritid" och "Universitetsinstitutioner". En mellanställning tycks sådana problem inta som gäller "Yrkesliv", "Anpassning", "Social utveckling", "Disciplin", "Psykoekologi" och "Socialisation". Denna relativt enkla analys styrker de förväntningar som kunde ställas på registrens innehåll. Ett studium av 40-registret i förhållande till de övriga ger de resultat som redovisas i tabellerna 2-7 på sidorna 33-36.

Register 40 innehåller 41 delfasetter. Dessa har studerats med avseende på de enskilda kategoriernas koincidenser med kategorierna i de övriga registren, 30, 33 + 34 ..., 70, 73 + 74 ..., 80, 83 ÷ 84 ... Som kriterium valdes att en 40-kategori måste koincidera 2 eller flera gånger med någon kategori ur dessa register.

I tabellerna 2-7 redovisas proportioner som är ett resultat av en radvis standardisering av de observerade frekvenserna.




Figur 1. Fördelning av fasetter

Tabell 2. Koincidenser mellan fasetter ur Problem-registret och Metod-registret

| | | Behovsanalys | Forskningsprogram | Dokumentation | Vetenskaplig diskussion | Historiska personer | Modeller | Erfarenheter | Forskningsmetodik | Handlingsprogram | Studie | Försök | Undersökning | Longitudinella studier | Utvärdering | Mätning | Enkät | Analys | Rapportering | Rekrytering | Studieträning | Betygsättning | Utveckling | Förslag | Prognos | Undervisning | Utbildning | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Register 30 | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 19 | 21 | 23 | 25 | 28 | 31 | 33 | 37 | 38 | 40 | 41 | Σ |
| Forsknings- och projektplanering | 1 | 14 | 14 | 14 | 28 | | | | | | 14 | | | | | | | | 14 | | | | | 14 | | | | 7 |
| Studieframgång | 2 | | | | | | 14 | | | | | | | | | | | 14 | | | | | | | 71 | | | 7 |
| Försök | 6 | | | | | | | | | | | | | | 73 | | 27 | | | | | | | | | | | 11 |
| Dataanalys | 7 | | | | | | | | 25 | | | | | | | | | 75 | | | | | | | | | | 4 |
| Mätinstrument | 8 | | | | | | | | | | | | | | | | | | 100 | | | | | | | | | 2 |
| Enkät | 9 | | | | | | | | | | | | | | | | | 33 | 67 | | | | | | | | | 3 |
| Modeller | 10 | | | | | | | | | | | | | | | | | | | | | | 100 | | | | | 3 |
| Resultatredovisning | 12 | | | 10 | | | | | 10 | | 10 | | | | 10 | 10 | | 10 | 40 | | | | | | | | | 10 |
| Elever | 14 | | | | | | | | 11 | 11 | 22 | | | 44 | | | | | 11 | | | | | | | | | 10 |
| Ungdom | 15 | | | | | | | | | | | | 100 | | | | | | | | | | | | | | | 2 |
| Forskningsprojekt | 16 | | | | | | | | | | | | | | | | | | 67 | | | | | 33 | | | | 3 |
| Pedagogisk debatt | 18 | | | | | 80 | | | | | | | | | | | | | | | | | | 20 | | | | 5 |
| Kognition | 24 | | | | | | | | | | 75 | | | | | | | 25 | | | | | | | | | | 4 |
| Lärarfunktion | 34 | | | | 10 | | | | 20 | | | | | | | 40 | | 10 | | | | | | | | | 20 | 10 |
| Språkfärdighet | 37 | | | | | | | | | | | | | | | 100 | | | | | | | | | | | | 7 |
| Studieteknik (material) | 39 | | | | | | | | | | | | | | | 67 | | | | | 33 | | | | | | | 3 |
| Undervisning (metodik) | 40 | | | | | | | | 38 | | 25 | 13 | | | 13 | | | | | | 13 | | | | | | | 8 |
| Material-metodsystem | 42 | | | | | | | | | | | | | | | 100 | | | | | | | | | | | | 4 |
| Klassen | 49 | | | | 25 | | | | | | | | | | 75 | | | | | | | | | | | | | 5 |
| Utredning/remiss | 50 | | | | | | | | 08 | | | | | | 33 | | | | 25 | | | 33 | | | | | | 3 |
| Perception | 53 | | | | | | | | | | 100 | | | | | | | | | | | | | | | | | 2 |
| Kreativitet | 56 | | | | | | | | | | 50 | | 50 | | | | | | | | | | | | | | | 4 |
| Attityder | 57 | | | | | | | | | | 33 | | 17 | | | | 17 | | | | | | 17 | | | 17 | | 6 |
| Personlighet (fasetter av) | 62 | | | | 33 | | | | | | 67 | | | | | | | | | | | | | | | | | 3 |
| Idrottsforskning | 67 | | | | | | | | | | | | | | | | 33 | 67 | | | | | | | | | | 6 |
| Frågor | 72 | | | | | | | | | | | | | | | 100 | | | | | | | | | | | | 2 |
| Fri skrivning | 89 | | | | | | | | | | | | | | | | | | | | | | 100 | | | | | 2 |
| Studie | 100 | | | | | | | 100 | | | | | | | | | | | | | | | | | | | | 7 |
| Formell bildning | 105 | | 33 | | | 67 | | | | | | | | | | | | | | | | | | | | | | 3 |

Tabell 3. Koincidenser mellan extensioner till Problemregistret och Metodregistret

| Register 33 | | Dokumentation | Vetenskaplig diskussion | Modeller | Forskningsmetodik | Handlingsprogram | Studie | Undersökning | Longitudinella studier | Utvärdering | Mätning | Enkät | Analys | Rapportering | Rekrytering | Användning | Skolprestationer | Betygsättning | Språkresa | Prognos | Utbildning | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 4 | 5 | 7 | 9 | 10 | 11 | 13 | 14 | 15 | 16 | 19 | 21 | 23 | 25 | 27 | 30 | 31 | 34 | 38 | 41 | Σ |
| Intelligens | 1 | | | | | | | | | | | | | 100 | | | | | | | | 3 |
| Personlighet (fasetter av) | 3 | | | | 25 | | | | | | | | 50 | 25 | | | | | | | | 12 |
| Stad | 6 | | | | | | | | | | | | | 100 | | | | | | | | 3 |
| Skolmiljö | 9 | 08 | | 08 | | | 16 | 16 | 08 | 33 | | | 08 | | | | | | | | | 12 |
| Samhälle | 13 | | | | | 33 | 67 | | | | | | | | | | | | | | | 5 |
| Avhandling | 15 | | | | | | | | | | 58 | | | | 08 | | | 08 | | 25 | | 12 |
| Skola | 17 | | | | 17 | | 33 | | | 17 | | | | 17 | | | | | 17 | | | 12 |
| Skolstadium | 18 | | | | | | 57 | 29 | | | | 14 | | | | | | | | | | 7 |
| Socialisationsteknik | 20 | | | | | | 09 | | | | 82 | | | | | | | 18 | | | | 11 |
| Undervisning (metodik) | 22 | | | | | | | | | | 100 | | | | | | | | | | | 2 |
| Inlärning | 25 | | | | | | | | | | 40 | | | | | | | 60 | | | | 5 |
| Språklig produktion | 29 | 100 | | | | | | | | | | | | | | | | | | | | 6 |
| Vuxna (med speciell funktion) | 30 | | | | | | | | 100 | | | | | | | | | | | | | 2 |
| Historisk forskning (enskilda begrepp) | 39 | | | | | | 57 | | | | | 29 | | | | | | | | | 14 | 7 |
| Lokal | 40 | | | | | | 67 | | | | | | | | | | 33 | | | | | 3 |
| Utbildningssystem | 46 | | | | | | | | | | 100 | | | | | | | | | | | 4 |
| Symposier | 52 | | | | 100 | | | | | | | | | | | | | | | | | 2 |
| Individuella differenser | 58 | 25 | 25 | | | | 50 | | | | | | | | | | | | | | | 3 |
| Muntlig framställning | 64 | | | | | | | | | | 100 | | | | | | | | | | | 3 |
| Prediktion | 79 | | | | | | | | | | 33 | | | 67 | | | | | | | | 4 |
| Kognitiva funktioner | 84 | | | | | | 75 | | | | | | | | | 25 | | | | | | 4 |
| Utveckling | 89 | | | | | 50 | | | | | | | 50 | | | | | | | | | 3 |
| Ungdom | 92 | | | | | | | | | | 100 | | | | | | | | | | | 2 |
| Prövning | 98 | | | | | | | | | | | | | | | | | | | 100 | | 2 |
| Högre studier | 99 | | | | | | | | | | 100 | | | | | | | | | | | 4 |

Tabell 4. Koincidenser mellan Instrumentregistret och Metodregistret

| Register 80 | | Vetenskaplig diskussion | Modeller | Erfarenheter | Handlingsprogram | Studie | Försök | Utvärdering | Mätning | Testkonstruktion | Enkät | Analys | Rapportering | Studieträning | Terapi | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5 | 7 | 8 | 10 | 11 | 12 | 15 | 16 | 18 | 19 | 21 | 23 | 28 | 35 | Σ |
| Nonsenstermer | 0 | | | | | 11 | 56 | 22 | 11 | | | | | | | 9 |
| Undervisning | 1 | | | 13 | 13 | | | | | 13 | | 13 | 13 | 38 | | 8 |
| Forskningsmetodik | 2 | 33 | | | | | | | | | | | | | 67 | 3 |
| Persongrupp | 3 | | | | | | 100 | | | | | | | | | 6 |
| Socialisationsteknik | 4 | 44 | | | | | 33 | 11 | | | 11 | | | | | 9 |

Tabell 5. Koincidenser mellan extensioner till Instrumentregistret och Metodregistret

| Register 83 | | Försök | Enkät | |
|---|---|---|---|---|
| | | 12 | 19 | Σ |
| Skola | 2 | 100 | | 3 |
| Lärarutbildning | 3 | 67 | 33 | 3 |
| Yrkesliv | 6 | 100 | | 3 |

Tabell 6. Koincidenser mellan Målregistret och Metodregistret

| Register 70 | | Disciplin | Forskningsprogram | Dokumentation | Vetenskaplig diskussion | Modeller | Erfarenheter | Forskningsmetodik | Försök | Utvärdering | Hypotesprövning | Testkonstruktion | Enkät | Problemsituationer | Analys | Observationsteknik | Statistisk skolning | Studieträning | Betygsättning | Skola | Undervisning | Utbildning | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 3 | 4 | 5 | 7 | 8 | 9 | 12 | 15 | 17 | 18 | 19 | 20 | 21 | 22 | 26 | 28 | 31 | 39 | 40 | 41 | Σ |
| Tidsbestämning | 2 | | | 33 | | | | | | | | | | | | | | | | 67 | | | 3 |
| Högre studier | 3 | | | | 17 | 17 | | | | | | | 50 | | | | | 17 | | | | | 6 |
| Skolform | 5 | 29 | 14 | 29 | | | | | | | | | | | | | | | | | | | 7 |
| Databehandling | 7 | | | | | 30 | | 30 | | | 20 | | | | 20 | | | | | | | | 10 |
| Kartläggning | 8 | | | | | | 08 | | | | | 08 | | 23 | | 23 | | 23 | | | | | 13 |
| Samarbete | 9 | | | | | | | | 100 | | | | | | | | | | | | | | 8 |
| Elever | 11 | | | | | | | | | 83 | | | | | | | | | | | | | 6 |
| Vuxna (med spec funktion) | 13 | 20 | | | | | | | | 07 | | 14 | | 07 | | | 40 | 07 | | | | 07 | 15 |

Tabell 7. Koincidenser mellan extensioner till Målregistret och Metodregistret

| Register 73 | | Erfarenheter | Försök | Utvärdering | Testkonstruktion | Problemsituationer | Betygsättning | |
|---|---|---|---|---|---|---|---|---|
| | | 8 | 12 | 15 | 18 | 20 | 31 | Σ |
| Nonsenstermer | 0 | 40 | | 40 | | | 20 | 5 |
| Yrkesliv | 2 | | | | | 20 | 40 | 5 |
| Vuxna | 8 | | 100 | | | | | 4 |
| Studerande | 9 | | 100 | | | | | 4 |
| Personal | 10 | | 100 | | | | | 8 |
| Lärarutbildning | 11 | | | | 100 | | | 2 |

En radvis och kolumnvis granskning visar påtagliga skillnader med avseende på vilka fasetter som koinciderar med varandra i de enskilda registren. En radvis granskning av tabell 2 visar t ex att "Språkfärdighet" i detta material endast koinciderar med "Mätning". Denna typ av information kan i en framtida återvinningsprocess användas för att kontrollera valet av ett "beteende" i ett datorprogram. Baseras proportionerna på en tillräckligt stor bas, kan de fungera som sannolikhetsvärden i en beslutstabell. Gäller informationssökarens intresse användningen av enkäter, kan det lätt konstateras att "Enkät" inom 30-registret leder till fasetterna "Idrottsforskning", "Försök" och "Attityder". Finns det mera specifik information i en sökfråga, så att register 33 kan användas, leder "Enkät" till fasett "Historisk forskning (enskilda begrepp)" och till "Skolstadium". Sökprocessen inom registret 80 leder till ännu mera specifik information nämligen till "Socialisationsteknik". Finns det dessutom en bestämning kan informationen preciseras ytterligare. Som framgår ur tabell 5 gäller det "Enkät" inom "Lärarutbildning".

Ingången "Enkät" visar med avseende på uttryckta mål att denna metod används för kartläggning av "Högre studier". Ingången "Enkät" saknas däremot i samband med målbestämningarna (se tabell 7).

Möjligheten att inkludera en beslutstabell på vilken ett val bland flera alternativa vägar kan byggas garanterar att en transformering, även om den är tillfällig, uppfattas av programmet som en process. Transformationer av den illustrerade typen är speciellt vanliga i samband med data som ett program får som input från en terminal, dvs när syftet är att åstadkomma ett användarorienterat system.

Att välja en väg så som illustrerats betraktas kanske ur konceptuell synvinkel som ett enkelt problem. Men detta problem har hittills lösts otillfredsställande bl a genom att Booleska operationer används. Dels tar det alltför lång tid att utföra dem, dels leder de till långa och därför svåröverskådliga uttryck.

## 6. AVSLUTANDE DISKUSSION

"Problem", "metod" och "mål" är grundkomponenter i forskningsprocessen. En vetenskaps kognitiva struktur antas bli kommunicerad genom vetenskapliga dokument. Att studera en titel till en vetenskaplig text med utgångspunkt i ett kognitionspsykologiskt perspektiv innebär ett studium av författares sätt att skapa en abstrakt representation av texten genom titelns påståendeform.

Den modell som ger uttryck för detta kallas i denna rapport PmM-modellen, som anger ett abstraherat påstående.

Titlar analyseras med utgångspunkt i "cues" som titelns manifesta struktur anger. Eftersom den manifesta utformningen av titeln kan vara fragmentarisk med avseende på PmM-paradigmet, har modellen försetts med "default"-variabler, som platshållare för saknade argument.

För att kunna nyttja relationer mellan titelns enskilda komponenter, behövs något sätt att ange (markera) komponenternas roller och strukturella samband. Kognitionspsykologisk forskning (Piaget & Inhelder, 1956; Miller & Johnson-Laird, 1976) tyder på att människans kognition nyttjar spatial organisering. Vuxna tycks bygga implicita kognitiva representationer (scheman) med hjälp av koordinater. Detta har på den lingvistiska nivån en motsvarighet i prepositionernas ordnande funktion.

Det fundamentalt annorlunda i att använda prepositioner som funktion vid en konceptuell analys är att den ambiguitet som de uttrycker på den konkreta nivån (i funktionen att relatera objekt till varandra) upphävs på den abstrakta nivån. Abstraktioner relateras i förhållande till deras intentioner eller extensioner, varvid sådana begrepp som lokalisering, riktning, punkt eller momentanitet får underordnad betydelse.

Några viktiga principer som styrt utvecklingen av programmet är:

1. Titlar består av en eller flera konceptualiseringar, varför avgränsningsregler är nödvändiga för att definiera gränserna för en konceptualisering.

2. Prepositioner pekar framåt, vilket betyder att när samtliga led efter varje huvudpreposition har bestämts återstår endast Metodkomponenten.
3. Titlar kan vara utan preposition eller endast innehålla underordnad preposition. Dessa blir alltid Problem, liksom titlar med final preposition, dvs som inte kan peka framåt.
4. Huvudkod får bara uppträda en gång per titel.

Datorprogrammet har prövats på en databas som är representativ för svensk utbildningsforskning. Basen bygger på 40 forskare, som genom obundet slumpvist stickprov valts ut ur populationen utbildningsforskare. I utprövningen ingår sammanlagt 660 verk och 8 865 referenser (se I. Bierschenk, 1979, kap 5).

Den bild som mönsterstudierna visar inom detta område kan verifieras (B. Bierschenk, 1974). Forskarna är i hög grad Problemcentrerade.

Utfallet av den funktionella kodningen har analyserats och sammanfattats i tabellform. Tabellerna ger en överblick över de problemområden som forskningen huvudsakligen har centrerats kring.

Genom en funktionell relatering av registren till varandra framgår tydligt i vilken utsträckning titlar innehåller information om tillämpad metod eller använda instrument. Det framgår också att mål sällan anges, vilket kan tyda på att de utsatta målen inte har uppnåtts när projektforskningens rapporter skrivs.

För generering av ett intermediärt språk krävs att koncepten som ligger till grund äger en för den intermediära nivån acceptabel abstraktionsgrad. Det betyder att vissa drag som karakteriserar naturligt språk inte kan tillåtas.

Ett nästa steg i utvecklingsarbetet är att identifiera koncept som inte befinner sig på den intermediära nivån och att skapa ordlistor för partiell mönsterigenkänning.

De register som skapats består av enheter i form av fraser. Eftersom utvecklingen av en återvinningsmekanism som ska fungera på basis av fraser är beroende av olika utprövningar är det för tidigt

att närmare analysera enheternas struktur. Däremot kan det sägas att de består av nominalkonstruktioner, främst ettords-koncept i obestämd form. Karakteristiskt är att begrepp med samma funktion, t ex att bestämma problemområde eller ange metodisk verksamhet, grupperas inom samma register. Denna generering av register bygger på "funktionell likhet", som är särskilt tydlig i Metod-registret. De nominal som finns i Problem-, Instrument- och Mål-registren är substantiv, t ex "Mätinstrument" (30), "Vuxna" (70) och "Intervjuer" (80). Den lexikaliska formen leder till föreställningen om ett "ting", vilket är lätt att acceptera inom dessa register. Det ingår också i den vanliga föreställningen att abstraktioner, t ex i form av verbalsubstantiv ("Testning" (40)) kan fungera som såväl ett "ting" som en "aktivitet". Men det förväntas inte att ord med form och innebörd som ger en konkret föreställning om ett ting skulle kunna ange handling, vara "verb". Enligt den kognitiva modellen som datorprogrammet nyttjar kan emellertid urskiljas och lyftas fram sådana dimensioner som en manuellt utförd analys troligen inte hade lyckats med på grund av dessa invanda föreställningar.

## 7. REFERENSER

Bierschenk, B. Perception, strukturering och precisering av pedagogiska och psykologiska forskningsproblem på pedagogiska institutioner i Sverige. Pedagogisk-psykologiska problem, Nr 254, 1974.

Bierschenk, B. A computer-based content analysis of interview texts: Numeric description and multivariate analysis. Didakometry, Nr 53, 1977.

Bierschenk, B. & Bierschenk, I. A system for a computer-based content analysis of interview data. (Studia Psychologica et Paedagogica, 32.) Lund: Gleerup, 1976.

Bierschenk, I. Försök med automatisk separering av referenser i en flerspråkig databas. Testkonstruktion och testdata, Nr 34, 1978.

Bierschenk, I. Intermediär språkstruktur. En metod för generering av ett språk för representation av vetenskaplig information. Stencil (Malmö: Pedagogisk-psykologiska institutionen), Juni, 1979.

Bunge, M. Scientific research I. The search for system. New York: Springer, 1967.

Miller, G.A. & Johnson-Laird, P.N. Language and perception. London: Cambridge University Press, 1976.

Newell, A. & Simon, H.A. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Piaget, J. & Inhelder, B. The child's conception of space. London: Routledge & Kegan Paul, 1956.

Wishart, D., Clustan 1C user manual. Stencil (London: Computer Centre, University College), 1975.

Department of
Educational and
Psychological Research
School of Education
Lund University
Malmö, Sweden

Abstract card

Reference card

Bierschenk, B., Bierschenk, I. & Sternerup-Hansson, A. Ett datorprogram för syntaktisk kodning av titlar till vetenskapliga skrifter. /A computer program for syntactic coding of titles to scientific documents./ <u>Testkonstruktion och testdata</u> (Malmö, Sweden: Dept. of Educ. and Psychol. Research), No. 36, 1979.

User oriented coding of knowledge that is mediated through titles of scientific papers requires information systems based on cognitive models. This report presents a computer program with comments for a conceptual coding of titles. The cognition model that governs program development is shortly described. The functioning of the program is illustrated by means of some computer outputs from a data base which is representative for educational research in Sweden. /In Swedish/

Indexed: Computer program, intermediate language, scientific documents.

Bierschenk, B., Bierschenk, I. & Sternerup-Hansson, A. Ett datorprogram för syntaktisk kodning av titlar till vetenskapliga skrifter. /A computer program for syntactic coding of titles to scientific documents./ <u>Testkonstruktion och testdata</u> (Malmö, Sweden: Dept. of Educ. and Psychol. Research), No. 36, 1979.